AMENDMENT TO THE CLAIMS

1-3. (Cancelled).

4. (Currently Amended) A computer readable medium having instructions, which when executed on a computer provide a user interface, the instructions comprising:

a speech synthesizer receiving input for synthesis and providing an audio output signal; and

a video rendering module receiving information related to the audio output signal, the video rendering module rendering a representation comprising a sequence of video frames of a talking head having a talking state with mouth movements in accordance with the audio output signal added to each of the frames during the talking state and a waiting state with added non-talking mouth movements during the waiting state in accordance with listening, and ~~The computer readable medium of claim 3~~ wherein the video rendering module returns to an earlier, preselected frame in the sequence upon reaching a selected frame in the sequence.

5. (Currently Amended) The computer readable medium of ~~claim 3~~claim 4 wherein the video rendering module tracks movements of the talking head in the sequence of video frames.

6. (Original) The computer readable medium of claim 5 wherein the video rendering module transforms affine parameters to physical movements of the talking head for each frame.

7. (Original) The computer readable medium of claim 6 wherein the physical movements include translations and rotations of the talking head.

8. (Original) The computer readable medium of claim 5 wherein the talking mouth positions are added based upon interpolated physical movements of the talking head.

9. (Original) The computer readable medium of claim 6 wherein for each of a plurality of frames, interpolated physical movements are calculated as a function of a corresponding preceding frame and a corresponding succeeding frame.

10. (Currently Amended) A computer readable medium having instructions, which when executed on a computer provide a user interface, the instructions comprising:

a speech synthesizer receiving input for synthesis and providing an audio output signal; and

a video rendering module receiving information related to the audio output signal, the video rendering module rendering a representation comprising a sequence of video frames of a talking head having a talking state with mouth movements in accordance with the audio output signal added to each of the frames during the talking state and a waiting state with added non-talking mouth movements during the waiting state in accordance with listening, wherein the video rendering module tracks movements of the talking head in the sequence of video frames, wherein the video rendering module transforms affine parameters to physical

movements of the talking head for each frame, wherein the physical movements include translations and rotations of the talking head and ~~The computer readable medium of claim 7~~ wherein for each of said plurality of frames, a mouth position corresponding to the talking state is added as a function of the physical parameters of the frame if a difference in at least one of physical parameters between the frame and the corresponding interpolated physical parameter exceeds a selected threshold, whereas if the difference in at least one of physical parameters between the frame and the corresponding interpolated physical parameter does not exceed the selected threshold, the mouth position corresponding to the talking state is added as a function of interpolated physical parameters.

11-15. (Cancelled).

16. (Currently Amended) A computer readable medium having instructions, which when executed on a computer provide a user interface, the instructions comprising:

a speech synthesizer receiving input for synthesis and providing an audio output signal; and

a video rendering module receiving information related to the audio output signal, the video rendering module rendering a representation of a talking head having a talking state with mouth movements in accordance with the audio output signal and a waiting state with mouth movements in accordance with listening, the video rendering module accessing a store having a sequence of

frames of the talking head and continuously rendering at least a portion of each of the frames in the sequence of frames while selectively adding a corresponding mouth position for the talking state to each of the frames in accordance with the audio output signal and in accordance with tracking movements of the talking head during the sequence of frames, wherein the video rendering module transforms affine parameters to physical movements of the talking head for each frame, wherein the physical movements include translations and rotations of the talking head, wherein the mouth positions are added based upon interpolated physical movements of the talking head, wherein for each of a plurality of frames, interpolated physical movements are calculated as a function of a corresponding preceding frame and a corresponding succeeding frame, and ~~The computer readable medium of claim 15~~ wherein for each of said plurality of frames, a mouth position corresponding to the talking state is added as a function of the physical parameters of the frame if a difference in at least one of physical parameters between the frame and the corresponding interpolated physical parameter exceeds a selected threshold, whereas if the difference in at least one of physical parameters between the frame and the corresponding interpolated physical parameter does not exceed the selected threshold, the mouth position corresponding to the talking state is added as a function of interpolated physical parameters.

17. (Currently Amended) A computer-implemented method for generating a talking head on a computer display to simulate a conversation, the method comprising:

continuously rendering a sequence of video frames of a talking head with each frame having mouth characteristics indicative of a non-talking state, wherein continuously rendering includes returning to an earlier, preselected frame in the sequence upon reaching a selected frame in the sequence;

tracking movements of the talking head throughout the sequence of video frames;

outputting a voice audio; and

selectively adding a corresponding mouth position to selected frames of the video sequence as a function of the voice audio and tracked movements of the talking head.

18. (Currently Amended) The computer-implemented method of ~~claim 17~~claim 23 wherein continuously rendering includes returning to an earlier, preselected frame in the sequence upon reaching a selected frame in the sequence.

19-22 (Cancelled).

23. (Currently Amended) A computer-implemented method for generating a talking head on a computer display to simulate a conversation, the method comprising:

continuously rendering a sequence of video frames of a talking head with each frame having mouth characteristics indicative of a non-talking state;

tracking physical movements including translations and rotations of the talking head throughout the sequence of video frames, wherein tracking movements includes transforming affine parameters to physical movements of the talking head for each frame;

calculating interpolated physical movements of the talking head as a function of a corresponding preceding frame and a corresponding succeeding frame for each of a plurality of frames

outputting a voice audio; and

selectively adding a corresponding mouth position to selected frames of the video sequence as a function of the voice audio and tracked movements of the talking head, and ~~The computer implemented method of claim 22~~ wherein adding a mouth position includes, for each of said plurality of frames, adding a mouth position corresponding to the talking state is added as a function of the physical parameters of the frame if a difference in at least one of physical parameters between the frame and the corresponding interpolated physical parameter exceeds a selected threshold, whereas if the difference in at least one of physical parameters between the frame and the corresponding interpolated physical parameter does not exceed the selected threshold, the mouth position corresponding to the talking state is added as a function of interpolated physical parameters.